Formulating multiple dependencies in sentence production

Shota Momma
Department of Linguistics, University of Massachusetts Amherst
Masaya Yoshida
Department of Linguistics, Northwestern University

Abstract

One of the defining properties of human language is the abundance of potentially unbounded dependencies between elements in a sentence. And yet, how speakers formulate dependencies in sentence production is still poorly understood. Here we examine the timing of verb planning in sentences involving two potentially different types of long-distance dependencies: *across-the-board* constructions (e.g., *Which book did you read and criticize?*) and *parasitic gap* constructions (e.g., *Which book did you read before criticizing?*). Using a new task we call the *Sentence-Word Interference* task, we show that speakers plan sentence-final verbs before sentence onset, but selectively when producing across-the-board sentences and not when producing parasitic gap sentences. When producing parasitic gap sentences, speakers plan sentence-final verbs relatively late, as they are uttering the pre-verb part of these sentences. Based on this timing contrast, we argue that speakers plan verbs before their dependents, but only when verbs and their pre-verbal dependents engage in both conceptual and direct syntactic (namely, selectional) relationships and not when they engage in only a conceptual relationship. We also argue that there are two distinct types of representations and processes for building across-the-board constructions (which involve coordinate structures) and parasitic-gap constructions (which involve subordinate structures) in sentence production. More broadly, the current study suggests that sentence planning is constrained by syntactic structures that are not reducible to conceptual structures and are richer than usually assumed in many prominent models of sentence production.

Correspondence concerning this article should be addressed to Shota Momma, Department of Linguistics, University of Massachusetts Amherst, N408 Integrative Learning Center, 650 North Pleasant Street Amherst, MA 01003. E-mail: snegishi@umass.edu

Introduction

Speaking often requires establishing potentially multiple, potentially unbounded dependencies, with limited memory resources in a limited amount of time. Given this challenge, the very fact that most adult native speakers can speak grammatically complex sentences reasonably well is impressive, and it suggests that speakers have sophisticated planning mechanisms that effectively coordinate multiple sub-processes that are necessary to produce a sentence. How exactly do speakers plan sentences, especially grammatically complex ones? Here we study the nature of sentence planning mechanisms by examining the timing of verb planning in the production of sentences where verbs and their dependents appear far apart. Based on this information about *when* speakers plan verbs, we aim to better understand *how* speakers establish dependency relationships between verbs and their dependents in the production of grammatically complex sentences.

In this article, we specifically focus on the production of sentences involving constructions known as the *Across-The-Board* (ATB) construction (De Vries, 2017; Ross, 1967; Williams, 1978) and the *Parasitic Gap* (PG) construction (Culicover & Postal, 2001; Engdahl, 1983; Ross, 1967), as in the following sentences (e represents phonologically null elements which are assumed to occupy the missing argument positions in some syntactic theories):

(1) a. Which book$_i$ did you read e$_i$ and criticize e$_i$? [ATB]

b. Which book$_i$ did you read e$_i$ before criticizing e$_i$? [PG]

These sentences involve multiple dependencies. The initial noun phrase filler (*which* NP) is interpreted in relation to both the first (*read*) and the second (*criticize*) verbs (i.e., *which* NP is interpreted as the theme argument of these two verbs). At what point in the production process do speakers plan the second verb (*criticize/criticizing*) in ATB and PG sentences? On the one hand, sentence production is often described as proceeding from "beginning-to-end" (Dell, 1986; Dell et al., 2008; V. S. Ferreira et al., 2018; V. S. Ferreira & Dell, 2000) with little need for look-ahead or advance planning of later-coming elements (Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; De Smedt, 1990; Kempen & Hoenkamp, 1987; Levelt, 1989). Given this consideration of incremental production, speakers may simply produce the filler without planning the verbs' lemma, and plan their lemmas only later, at the point when they need to articulate them. On the other hand, verbs often encode grammatically critical information about their dependents. So speakers may have to plan verbs early, before their dependents can be grammatically encoded and thus articulated. In the case of ATB and PG constructions, the second verb bears conceptual and grammatical relationships to the initial element *which book*, so speakers may plan the second verb before the articulation of the sentence initial filler. This question of timing is important, because *when* speakers plan verbs gives us a clue about *how* speakers establish dependencies between elements in a sentence. To better understand how speakers establish dependencies, we must first understand which piece of a sentence is constructed at which point in time during sentence production. The primary goal of the current study is to examine the timing of verb planning in sentences where the initial filler forms dependencies with multiple verbs. As we will see below, information about when speakers plan verb in

ATB and PG sentences has rich implications for theories of sentence production as well as theories of grammatical representations.

The main argument we will make is that speakers (tend to) plan verbs before their syntactic complement but not necessarily before the noun phrases denoting the patient/theme of the relevant verbs, and that ATB and PG constructions involve distinct types of syntactic dependencies and correspondingly distinct types of planning procedures. More broadly, we will suggest that sentence production is guided by syntactic dependencies that are not reducible to conceptual relationships.

The timing of verb planning in sentence production

Verbs carry information that are crucial to determining the semantic, syntactic, and sometimes morpho-phonological properties of other elements of sentences. So, to speak sentences grammatically, verbs' lemmas (cf. Kempen and Huijbers, 1983; Levelt et al., 1999) may be necessary when the grammatical properties of pre-verbal elements in some way critically depends on information encoded in them. Accordingly, some theories of sentence production assume a critical role for verbs in grammatical encoding, suggesting that verbs' lemmas may be planned early in sentence planning processes (e.g., Bock and Levelt, 1994; F. Ferreira, 2000).

Various studies have investigated the timing of verb planning in the production of simple sentences (Hwang and Kaiser, 2014; Kempen and Huijbers, 1983; Lindsley, 1975, 1976; Momma and Ferreira, 2019; Momma et al., 2016, 2018; Sauppe, 2017; Van de Velde et al., 2015, see also Griffin and Ferreira, 2006). Many of the studies investigating the timing of verb planning use a methodology known as extended picture-word interference (ePWI, Damian and Dumay, 2007; Hwang and Kaiser, 2014; Meyer, 1996; Miozzo and Caramazza, 1999; Momma and Ferreira, 2019; Momma et al., 2016, 2018; Schriefers and Teruel, 2000; Schriefers et al., 1998) ePWI is an extension of the typical picture-word interference task (Lupker, 1979; Schriefers et al., 1990 *inter alia*). In a simple PWI study, speakers name a picture while ignoring distractor words that are visually or auditorily presented. A well-established effect in simple PWI experiments is the *semantic interference effect*, a delay in speech onset due to distractors that are semantically and categorically related to the target word (e.g., a distractor *dog* delays the onset of speaking *cat*) (Lupker, 1979; Schriefers et al., 1990 among others; see Bürki et al., 2020 for a recent meta-analysis). ePWI is an extension of PWI, where speakers produce multi-word utterances while ignoring distractors (Damian & Dumay, 2007; Meyer, 1996; Momma & Ferreira, 2019; Momma et al., 2016, 2018; Schriefers et al., 1998). The semantic interference effect in the ePWI task can be used to examine the timing at which a particular lemma is planned in phrase-level or sentence-level production. The logic is that, if the semantic interference effect is observed in the onset latency of utterances, speakers must have already initiated the selection process for the word being interfered with before they start speaking. The locus of the semantic interference effect is suggested to be at the level of lemma selection (Damian & Bowers, 2003; Hwang & Kaiser, 2014; Jescheniak et al., 2020; Lupker & Katz, 1981; Schriefers et al., 1990). Thus, to the extent that speakers experience a semantic interference effect in speech onset latency, it can be inferred that speakers predominantly plan the lemma of the interfered word before starting to speak. ePWI can further be extended by measuring not only the latency of speech onset, but also the production time (defined as the time

interval between the onset of a unit and the onset of the next unit) of a particular part of a sentence (Momma & Ferreira, 2019; Momma et al., 2018). If speakers take more or less time saying a particular part of a sentence given a semantically related distractor, they must be performing the lemma selection of the interfered word at that time. Thus, ePWI offers an useful tool for studying the time-course of lexical planning in sentence production.

ePWI has been used to study the timing of verb planning across different languages (Korean: Hwang and Kaiser, 2014; Japanese: Momma et al., 2016; German: Schriefers et al., 1990; English: Hwang and Kaiser, 2014; Momma and Ferreira, 2019; Momma et al., 2018) and different constructions (active transitive: Hwang and Kaiser, 2014; Schriefers et al., 1990; unergative: Momma et al., 2018; Schriefers et al., 1998; short and long passives: Momma, 2016; unaccusatives: Momma and Ferreira, 2019; Momma et al., 2018). We will not extensively review the details of those studies here, but across studies, the semantic interference effect in speech onset latency, that is, the sign of advance verb planning before the sentence onset, was consistently absent when the sentence-initial element was agentive (as in SOV transitive and unergative SV sentences in English and German Momma, 2016; Momma and Ferreira, 2019; Momma et al., 2016, 2018; Schriefers et al., 1998; cf. Hwang and Kaiser, 2014 but see also Momma et al., 2016 for a discussion). In comparison, it was consistently present when the sentence-initial element was non-agentive (mostly patient or theme, as in object-initial sentences in Japanese and passive and unaccusative sentences in English; Momma and Ferreira, 2019; Momma et al., 2016, 2018). Thus, the results of various ePWI studies on the timing of verb planning could be captured by the following generalization.

- Speakers tend to plan verbs before the production of their underlying objects but not their underlying subjects.

Note that the term *underlying object* includes the subjects of passive and unaccusative verbs, as well as canonical objects of usual transitive sentences. This generalization is natural given the closer tie between verbs and their non-agentive arguments, as assumed in many modern linguistic theories (see especially: Baker, 2009; Harley and Stone, 1992; Kratzer, 1996, 2003; Marantz, 1981). But this generalization could be stated in two different ways. On the one hand, it could be stated in terms of conceptual or semantic relationships between patient/theme roles and events denoted by verbs. Under this *conceptual account*, speakers plan verbs before the production of the constituent bearing the theme/patient role of the event denoted by those verbs, because theme/patient roles are dependent on verbs in some critical way (a view that is adopted by e.g., Kratzer, 2003). Under this account, the generalization can be reformulated as follows:

- Speakers tend to plan verbs before the production of a constituent bearing their theme/patient role but not the constituent bearing the agent role. [Conceptual account]

On the other hand, the generalization can be stated in syntactic terms. Under this *syntactic account*, verbs are more likely to be planned before the production of their (underlying) syntactic complement, because verbs syntactically select (that is, subcategorize) their complements. Under this account, the generalization can be reformulated as follows:

- Speakers tend to plan verbs before the production of their syntactic complement(s) but not their subjects. [Syntactic account]

The conceptual and syntactic accounts make overlapping predictions on the timing of verb planning in most cases. In the vast majority of sentences, the constituent receiving a verb's theme/patient role is also that verb's syntactic complements underlyingly. But as we explain in the next section, at least under some theories of syntax, there are types of sentences where verbs' semantic objects are not their syntactic complements. Those sentences involve the construction known as the *parasitic gap* construction, which we will compare to a superficially similar construction, known as the *across-the-board* construction.

Across-the-board and parasitic gap constructions

In English wh-movement constructions, one filler is usually associated with one gap. However, there are cases in which a filler is associated with multiple gaps in a class of constructions known as multiple gap constructions (Levine & Sag, 2003). In English, there are roughly two types of multiple gap constructions. One is observed typically in a coordination context, namely ATB extractions like (1a) (Ross, 1967; Williams, 1978), and the other is observed typically in sentences with clausal adverbial adjuncts, namely PG constructions like (1b) (Engdahl, 1983; Ross, 1967). ATB and PG constructions are very similar. Both involve one filler associated with two verbs (and thus two gaps), as we can see in (1a) and (1b). Furthermore, if the first gap is replaced with an overt noun phrase, these constructions are unacceptable as we can see in (2a) and (2b).

(2) a. * Which book$_i$ did you read it and criticize e$_i$? [ATB]
b. * Which book$_i$ did you read it before criticizing e$_i$? [PG]

Furthermore, ATB and PG constructions can express similar meanings. For example, (1a) and (1b) express similar meanings with similar temporal relations between two events. In both sentences, the most natural reading is that the event described in the first clause precedes (takes place prior to) the event described in the second clause. Note, in a coordination construction like (1a), this interpretation is not obligatory (i.e., a simultaneous interpretation is possible, in which the two events described by two conjuncts take place simultaneously), but such an interpretation is most salient in an example like (1a).

There are also systematic differences between ATB and PG constructions. In general, the distribution of PG constructions is more restricted than that of ATB constructions (see Postal, 1993 for an extensive comparison between PG and ATB constructions). For example, PG constructions are less acceptable inside a tensed adjunct clause (Culicover, 2001; Engdahl, 1983). A PG cannot be licensed by a certain type of operation known as *A-movement*, which is involved in the derivation of passive and raising sentences (among other kinds of sentences) (Culicover, 2001; Engdahl, 1983). The implicit subject of the second clause in PG constructions must be bound by the subject of the clause containing the gap that hosts the PG (Williams, 1992). ATB and PG constructions show different profiles in terms of pronominal binding (cf. *reconstruction effects*: Haik, 1985; Munn, 1992; Nissenbaum, 2000; but see Bruening and Al Khalaf, 2017). Perhaps the most striking difference between ATB and PG is the distribution of the second gap. In PG constructions,

the second gap contained in the adjunct clause is optional; thus it can be replaced by a pronoun or an overt noun phrase. In comparison, the second gap in ATB cannot be replaced by an overt noun phrase, as shown below.[1]

(3) a. * Which book$_i$ did you read e$_i$ and criticize it$_i$/the author? [ATB]
b. Which book$_i$ did you read e$_i$ before criticizing it$_i$/the author? [PG]

Given these similarities and differences, there are roughly speaking two analytical approaches to ATB and PG constructions. One approach posits that ATB and PG constructions are structurally alike (Bruening & Al Khalaf, 2017; Chaves, 2012; Hornstein & Nunes, 2002; Levine & Hukari, 2009; Munn, 1992; Sag, 1983; Williams, 1990). We call this approach the *identity approach*, in the sense that ATB and PG constructions are analyzed as involving the same type of syntactic dependencies. Under the identity approach, the two gaps in both ATB and PG sentences are generated by the same syntactic operations and share the same syntactic properties, that is, they are sensitive to the same syntactic constraints. A competing approach posits that ATB constructions are structurally different from PG constructions (Chomsky, 1986; Contreras, 1984, 1993; Haik, 1985; Nissenbaum, 2000; Postal, 1993). We call this approach the *non-identity approach*. Typically, under the non-identity approach, a phonetically empty element (a null operator, represented as *Op*), is assumed to be the syntactic complement of the second verb underlyingly (Browning, 1982; Chomsky, 1986; Contreras, 1984, 1993). This null element is co-referential with the wh-filler, and as a consequence, the filler is interpreted as the theme of the second verb. The following spells out the relevant syntactic difference assumed in the non-identity approach.

(4) a. Which book$_i$ did you read t$_i$ and criticize t$_i$? [ATB]
b. Which book$_i$ did you read t$_i$ before Op$_i$ criticizing t$_i$? [PG]

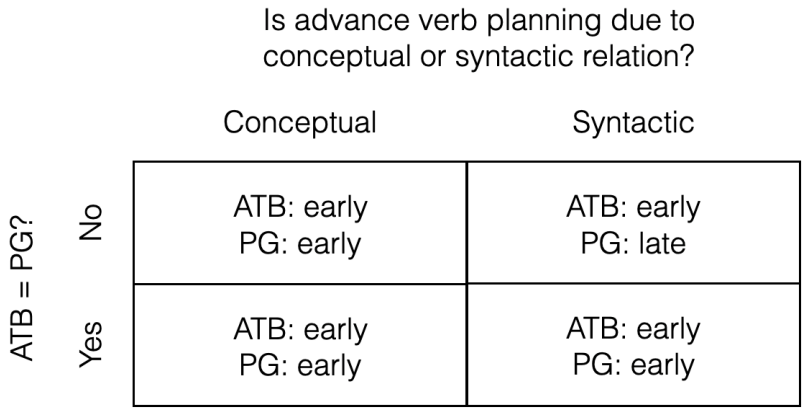
Most relevantly, under this non-identity approach, the filler is not the direct syntactic complement of the second verb in the PG construction. In other words, the filler is not directly linked to the second verb. This is because a null element (Op) mediates the relationship between the filler and the second verb in PG sentences, unlike in ATB sentences. In comparison, in ATB sentences, the filler is underlyingly the genuine syntactic complement of the second verb. In other words, the filler is directly linked to the second verb. Crucially, regardless of which approach one adopts, the conceptual relationship between the filler and the second verb in minimally different ATB and PG sentences is identical. For example, the filler (*which book*) is interpreted as the theme of *criticize/criticizing* in both (1a) and (1b). Thus, only under the non-identity approach are the syntactic and conceptual relationships between the filler and the second verb dissociated in PG constructions. Therefore, by comparing the production of ATB and PG sentences, we can examine whether the conceptual relationship (between verbs and the constituents that saturate their theme/patient-role) or the syntactic relationship (between verbs and their syntactic complements) is responsible for triggering advance planning of the second verb.

[1]Similar cases where one of the gaps is missing in ATB constructions have been observed (Goldsmith, 1985; Kehler, 2002; Ross, 1967; Schmerling, 1972, *inter alia*). However, in many of these cases, the gap position does not exist, rather than it being replaced by an overt NP. Thus, something different is going on in these cases. Though we recognize these exceptions in ATB constructions, we contend that they are different from cases like PG constructions where the second gap is replaced by an overt NP.

Hypotheses and predictions

The two hypotheses for the timing of verb planning (conceptual vs. syntactic accounts) and the two hypotheses about the structures of ATB and PG (identity vs. non-identity accounts) combine to create four different hypotheses about how speakers might plan ATB and PG sentences. Figure 1 summarizes these four hypotheses and associated predictions regarding the timing of verb planning in ATB and PG sentences. In Figure 1, *early* means before the articulation onset of the object wh-filler, while *late* means after the articulation of the wh-filler and before the verb is actually spoken. Only one of the four hypotheses makes a diverging prediction. Namely, only under the combination of the syntactic account and the non-identity account is it predicted that speakers plan verbs before uttering the filler in ATB sentences but not in PG sentences. The primary goal of the current study is to test this prediction.

Figure 1. Predicted timing of the planning of the second verb in ATB and PG sentences under four different hypotheses.



It is worth noting here that this critical prediction is difficult to derive from differences orthogonal to the structural differences posited by the non-identity account. Certainly, ATB and PG are different in multiple respects as briefly discussed above. However, some documented differences between ATB and PG actually lead to precisely the opposite prediction, as they suggest that the relationship between the first and the second clauses are, if anything, *tighter* in PG constructions than in ATB constructions. For example, the presence of the adverbial phrase that hosts the second gap in the PG construction is dependent on the presence of the main clause (i.e., an adverbial clause cannot exist without a main clause), but conjuncts in a coordinate structure can contain an independent full-fledged clause. And the adverbial clause modifies the event described in the main clause but a conjunct in the coordinated structure is not necessarily a modifier of the event described in the other conjunct, i.e., the interpretation of the adverbial clause is determined in relation to the event described by the main clause but such is not always the case for conjuncts in a coordinate structure. Furthermore, the participle adjunct clause that hosts a PG includes an implicit subject, the interpretation of which is dependent on the matrix subject (Williams, 1992),

but ATB constructions do not include such an implicit subject.

These differences between PG and ATB suggest that the interpretive relationship between the main and non-main clauses in PG constructions is stronger in PG than in ATB constructions. Furthermore, the tense of the non-main clause in PG constructions, but not necessarily in ATB constructions, must be interpreted in relation to the tense of the main clause (Ogihara, 1994), suggesting again that the relationship between the main and the non-main clauses in PG is stronger in PG constructions than in ATB constructions. Because these observations suggest that the non-main clause has a tighter relationship to the main clause in PG than in ATB, it could be predicted that, if anything, the verb of the non-main clause should be planned earlier in PG than in ATB. This is contrary to the prediction derived above. Thus, although it is true that ATB and PG can differ in more than one way, it is unclear how differences other than the syntactic difference postulated by the non-identity account could derive the critical prediction.

Sentence-Word Interference (SWI) paradigm

One of the methodological challenges in investigating the production of complex sentences is that it is difficult to elicit complex sentences reliably. To our knowledge, there is no clear way of eliciting sentences containing ATB or PG constructions using picture stimuli, which means that the extended picture-word interference paradigm is not easily applicable to studying the timing of verb planning in ATB and PG sentences. Thus, to study the time-course of planning processes involved in the production of ATB and PG sentences, we must first develop a task that allows the reliable elicitation of those types of sentences.

In the current study, we developed a variant of a Rapid-Serial-Visual-Presentation (RSVP)-based sentence recall task. In this task, participants read target sentences in the RSVP fashion, and recall those sentences after some secondary tasks. As a working hypothesis, we adopt the *regeneration hypothesis* (Lombardi & Potter, 1992; Potter & Lombardi, 1990, 1998). The regeneration hypothesis maintains that people can remember sentences well not because they form a verbatim memory of sentences but because they "regenerate" the target sentence from conceptual memory using the (un-ordered) set of recently activated words. In support of this hypothesis, Potter and Lombardi, 1990 showed that, when recalling a sentence, speakers often substitute words in a sentence with recently activated lure words as frequently as 34% of the times (see also Baddeley et al., 2009; Rummer et al., 2013; Sachs, 1974). Under the regeneration hypothesis, the process of recalling a sentence involves the process of sentence production (see Bock, 1982 for a similar claim), that is, it involves the processes of mapping conceptual representations onto syntactic representations, and syntactic representations onto phonological representations (Bock & Levelt, 1994; Garrett, 1975; Levelt, 1989).

The sentence recall task allows us to elicit across-the-board and parasitic gap sentences relatively reliably. But how can we study the timing of verb planning using the sentence recall task? Here we extended the recall task based on the same logic as the ePWI task. The idea is to present a distractor verb that is sometimes semantically related to the verb in question right when speakers are prompted to recall sentences. If we can establish that speakers indeed slow down due to distractor verbs in sentence recall (as we will test in Experiment 1 below), the interference effect in a recall task can be used as a tool for investigating the timing of verb planning in sentence regeneration, which by hypothesis

reflects the processes of sentence production. The detailed structure of this *Sentence-Word interference* task is shown in Figure 2 and described in the Procedure section below.

Method

Participants

Forty-eight (Experiment 1) and one hundred and fifty-five (Experiment 2) native monolingual English speakers were recruited using Prolific Academic. Experiment 2 had a larger sample size than Experiment 1 because we expected that Experiment 2 would have greater amount of noise in the data than Experiment 1 due to the increased complexity of sentences (which would lead to less accurate recall), and because Experiment 2 tested our critical predictions. For each participant, informed consent was obtained. Each experiment took roughly between 20–30 minutes. Each participant was paid five US dollars for compensation. Those participants who did not follow instructions or whose recordings were too poor in quality (four participants in Experiment 1 and fifteen participants in Experiment 2) or had fewer than three (i.e., <25%) error-free trials in any of the conditions (two participants in Experiment 1 and twenty nine participants in Experiment 2) were replaced.

Materials

Sixty-four sentence frames like the following were constructed for Experiment 1.

(5) a. Which computer did the engineer test? [First verb]
b. Which computer did the engineer repair? [Second verb]

The first sentence (5a) was in the *first verb* conditions, and (5b) was in the *second verb* conditions, because the verb *test* and *repair* were used as the first and the second verb of ATB and PG sentences in Experiment 2 (which involved sentences like *Which computer did the engineer test and repair/before repairing*). Those sentences always had 'which NP' as the filler associated with the object position of a transitive verb. The subject NP varied across sentences in Experiment 1 (but not in Experiment 2, see below). For each sentence frame, we chose a distractor verb that is related to the verb that was used as the second verb in Experiment 2, as in (5b) (e.g., *help* for *repair*). Distractor verbs were chosen on an intuitive basis, but their relatedness to the target verbs was later verified using Latent Semantic Analysis (Landauer & Dumais, 1997). Note that all distractor verbs were also used as target verbs in some other trials in the second verb condition condition. This decision was based on the observation that, in a Picture-Word interference task, distractors that are also in the response set are more likely to elicit a larger semantic interference effect (Roelofs, 1992; cf. Caramazza and Costa, 2000). Although the current task is not the same as a PWI task and participants in the current task could not know what verbs were in the response set (unlike in a PWI task with picture familiarization), it is possible that distractor verbs that were used as target verbs in preceding trials could increase the magnitude and/or likelihood of observing a semantic interference effect. Because the logic of Experiment 2 relies on a reliable semantic interference effect from verb distractors, we wanted to maximize the chance of obtaining a semantic interference effect.

In Experiment 2, using exactly the same set of filler NPs and verbs as in Experiment 1, sixty-four sentence frames like the following were constructed.

(6) a. Which computer did you test and repair? [ATB]

b. Which computer did you test before repairing? [PG]

The adverb introducing the adverbial clause in PG sentences was always *before*. The reason for this decision despite the availability of other adverbs (such as *after*, *while*, etc.) was twofold. First, we wanted to hold the number of function words used in the ATB and PG conditions constant at one. Because *and* was practically the only connective that could be used in relatively natural and easy-to-remember ATB sentences (words like *or* and *but* create pragmatically odd sentences unless other licensing contexts are given), we also wanted to use only one type of adverb introducing PG clauses. Second, the ATB sentences with *and* and PG sentences with *before* are minimally different in their meanings. The most natural interpretations of (6a) and (6b) are very similar, in that they both involve two past events that occurred in succession, with the event described by the main clause verb preceding the event described by the verb in the adverbial clause. If other adverbs, such as *after* and *while* were used, the differences in meaning would increase between conditions, introducing additional confounds.

Also, the subject was always *you* in Experiment 2, although in Experiment 1, the subject varied across sentences. This difference was due to our pilot study for Experiment 2, which was conducted after Experiment 1. In Experiment 2, the critical sentences were more complex than in Experiment 1, and according to the pilot study for Experiment 2, this made the sentence recall task too hard for many participants. To lessen the difficulty of memorizing and recalling sentences in Experiment 2, we made the decision to simplify sentences by using only *you* as the subject. The distractor verbs and their pairing with target sentences were identical to Experiment 1. There were no filler sentences in Experiments 1 or 2, because based on our pilot experiments using similar tasks, including sentences with various different structures reduced the chance of speakers producing the intended sentences, thereby adding noise to the data.

Because Experiment 2 was the critical test of our main hypotheses, we measured the acceptability as well as plausibility of the target sentences used in Experiment 2, based on two separate norming experiments each with twenty-four participants. The results of these norming experiments are summarized in Table 1, along with the acceptability and plausibility scores of filler sentences, which comprised sentences with various degrees of acceptability and plausibility. As can be seen in Table 1, there was a slight advantage for the ATB sentences both in terms of acceptability and plausibility. A linear mixed effects model of the z-score transformed acceptability score with Sentence Type (ATB vs. PG) as fixed effects with a maximal random effects structure revealed that this ATB advantage was indeed significant ($\hat{\beta} = -0.072$, $SE = 0.032$, $p = .027$). An analogous effect was found on plausibility scores ($\hat{\beta} = -0.113$, $SE = 0.046$, $p = .019$). However, because the difference was slight (a 0.15 point difference in acceptability and a 0.25 point difference in plausibility), and because it was not easy to match those scores across conditions while simultaneously satisfying various other constraints, we decided to statistically control for the effect of plausibility and acceptability, instead of attempting to create perfectly matched stimuli. Thus, in the statistical analyses in Experiment 2, we initially included the plausibility and acceptability scores as well as their interaction with Relatedness as covariates in all models. However, in all of the models reported here, there was no evidence that plausibility, acceptability, the

interaction between plausibility and Relatedness, or the interaction between acceptability and Relatedness affected onset latency or production time. Including those variables also did not change the results. Therefore, we did not include these variables in our final model and we only report the results of the models without them.

Table 1
Mean acceptability and plausibility scores of target sentences in Experiment 2.

Sentence Type	Acceptability	Plausibility
ATB	6.17 (0.44)	5.98 (0.28)
PG	6.02 (0.37)	5.74 (0.17)
Good filler	6.33 (0.54)	6.51 (0.55)
Bad filler	1.91 (-1.57)	2.33 (-1.53)

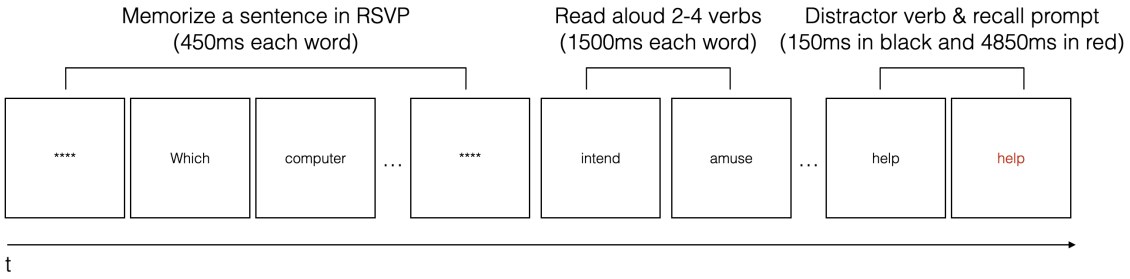
Procedure

Experiment 1 and Experiment 2 had identical procedures. They were run online using PCIbex (Zehr and Schwarz, 2018). In both experiments, participants were first given an explanation of the task. The instructions given were as follows.

> *In this experiment, you will first read a sentence in a word-by-word fashion. Each word will be flashed on the screen. Your task is to read the sentence silently and memorize the sentence for later recall. After you see each sentence, you will see a series of words. Your task is to read aloud each word as it is presented to you as soon as possible, but EXCEPT when the word is presented in red font. Note that words can be briefly in black font before turning red. When you see a word in red font, INSTEAD OF reading the word aloud, you should recite the sentence you memorized aloud, as soon as possible.*

Following the instructions, participants went through four practice trials, which had the same task structure as the experimental trials. Following the practice trials, the main experimental trials began, which had the following sequence of events. First, a string of four asterisks (****) was presented for 450 milliseconds, followed by an RSVP presentation of a sentence with 450 milliseconds/word presentation rate. The sentence was followed by four asterisks, which marked the end of each sentence. Following the sentence to be memorized, a series of pseudo-randomly chosen verbs (which had no obvious phonological or semantic relationship to the preceding sentence) were presented one after the other, and participants were instructed to read aloud each word as they saw it. There were 2–4 words in this sequence, and each word was presented for 1500ms. After the last word, the distractor word that was either semantically related or unrelated to the relevant verb in the memorized sentence was presented, initially in a black font (which made it indistinguishable from the preceding series of words). The distractor word's font color then changed to red, at which point speakers recalled the sentence instead of reading aloud the distractor verb. The distractor verb remained on the screen for 4850 milliseconds after it turned red, and during this time speakers' recall responses were recorded. Each trial was separated by a

Figure 2. A schematic illustration of the task in Experiments 1 and 2.



prompt to press the space key for the next trial. The distractor verb was initially presented in a black font for 150 milliseconds to increase the likelihood that speakers would register the distractor word at least partially. A schematic illustration of the task is shown in Figure 2.

Scoring and analyses

All audio files were first transcribed and coded for errors. Errors were defined as any deviation from the target sentence, with the following exceptions. First, substitutions of non-critical words with synonyms were tolerated (e.g., *warm up* substituted with a synonym *heat up*). In addition, in ATB sentences in Experiment 2, speakers sometimes said *and then* instead of just *and*, but those trials were not counted as errors. Incomplete utterances, trials where participants read aloud the distractor word, trials where participants were still uttering the previous word after the recall prompt, and trials where participants uttered overt hesitations (uh, um, etc.) before finishing the sentence were also coded as erroneous.

In both Experiments 1 and 2, the onset latency of the error-free utterances were manually extracted using Praat, by the author who was blind to the relatedness condition. Additionally, in Experiment 2, we aligned the audio of the error-free trials with their transcriptions using an automatic text-to-speech forced aligner called *Gentle*, followed by human adjustments for trials where the forced aligner suggested implausibly short durations. The alignment data was used to extract the production time of each word, defined as the time interval between the onset of a word and the onset of the next word. The production time for each word in Experiment 1 was not calculated, because it is not relevant to our theories and predictions. In Experiment 1, the dependent variable of interest was onset latency. In Experiment 2, in addition to the onset latency, we analyzed the production time of the word preceding the target verb, that is, the time speakers spent on the *before/and* region. This region was examined to verify that the verb interference effect was indeed present in PG sentences, just at a different timing than in the ATB sentences. This region was chosen *a priori* as the region of interest, based on our previous (unpublished) ePWI study that showed the verb interference effect in the production time of the word immediately preceding the target verb. The onset latency measure should reflect the process that occurs prior to sentence onset, and the production time of the pre-verbal region should reflect the process that happens right before the production of the relevant verb. Of course, finding

the verb interference effect in this region does not mean that speakers only plan verbs right before articulating them, but that would at least show that the semantic interference effect is present in PG sentences, just not in the onset latency.

All analyses were conducted in R (R Core Team, 2020), using the *brms* package (Bürkner, 2018). For both onset latency and word-by-word production time, we fit Bayesian hierarchical models, with the maximal random effects structure in the sense of Barr et al. 2013. In Experiment 1, the model included Verb Position (first vs. second, coded as -0.5 and 0.5), Relatedness (related vs. unrelated, coded as -0.5 and 0.5), and their interaction as fixed effects; by-subject and by-item random intercepts; and by-subject and by-item random slopes for Verb Position, Relatedness, and their interaction. In Experiment 2, for onset latency, the model included Sentence Type (ATB vs. PG, coded as -0.5 and 0.5), Relatedness and their interaction as fixed effects; by-subject and by-item random intercepts; and by-subject and by-item random slopes for Sentence Type, Relatedness and their interaction. The same model structures were used for the production time analysis in the pre-verbal region. In both experiments, we set normal priors over all fixed effects and intercepts. All priors had a mean value of 0, and the variance on the prior distribution was set to 1 for all fixed effects, and 10 for the intercepts. The prior on the correlation matrix was a LKJ regularizing prior (Lewandowski et al., 2009) with $\eta = 2$ (Vasishth et al., 2018). For the production time analysis, production times of more than 1000ms were removed (64 out of 9929 data points). For all analyses, we used 95% credible interval (based on percentile) to make inferences. That is, we conclude that there is evidence for an effect if the 95% credible interval for that effect does not include 0. For each model, four Monte Carlo Markov Chains with 5000 samples were run. The first 2500 samples were discarded as a warm-up period. For all models reported below, the $\hat{R}$ statistic was at or near 1.0 for all fixed effects parameters of interest, and no divergences were observed.

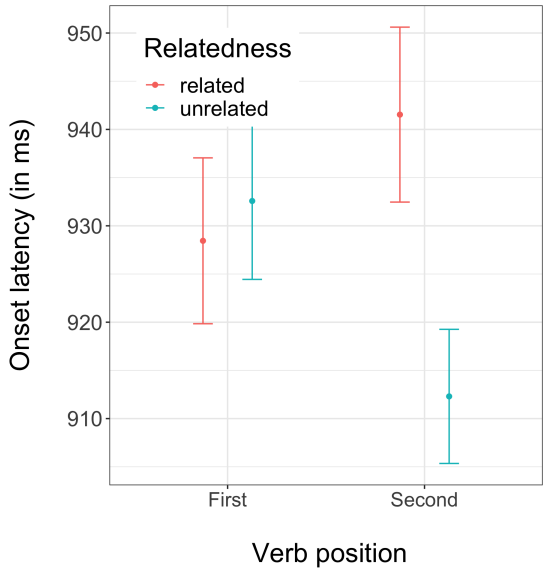
Results

Experiment 1

Overall, 20.3% of trials (636 out of 3136 trials) were identified as erroneous and were excluded from subsequent analyses. The error rates across the four conditions were as follows: First verb-Related: 20.2%; First verb-Unrelated: 17.6%; Second verb-Related: 22.5%; Second verb-Unrelated: 20.2%. One of the items (item #10 in the appendix) had zero error-free trials in the Second verb-Related condition so it was excluded from the subsequent onset latency and production time analyses.

As can be seen in Figure 3, in the first verb condition, if anything, speakers were on average 6 milliseconds *faster* to start speaking given a related distractor than given an unrelated distractor, suggesting that related distractors did not interfere with the verbs that were used as the first verb in Experiment 2. In contrast, speakers were on average 24 milliseconds slower to start speaking in the second verb condition given a related distractor than given an unrelated distractor, suggesting that related distractor words interfered with the verbs that used as the second verb in Experiment 2. Supporting this pattern, there was evidence for an interaction between Relatedness and Verb Position ($\hat{\beta} = -0.040$, 95% CrI $= [-0.078, -0.003]$, $Pr(\beta < 0) = 0.995$). There was no clear evidence for a main effect of Relatedness ($\hat{\beta} = -0.014$, 95% CrI $= [-0.033, 0.005]$, $Pr(\beta < 0) = 0.943$) or evidence for a

Figure 3. By-subject mean onset latency across four conditions in Experiment 1. Error bars represent standard error of the mean.



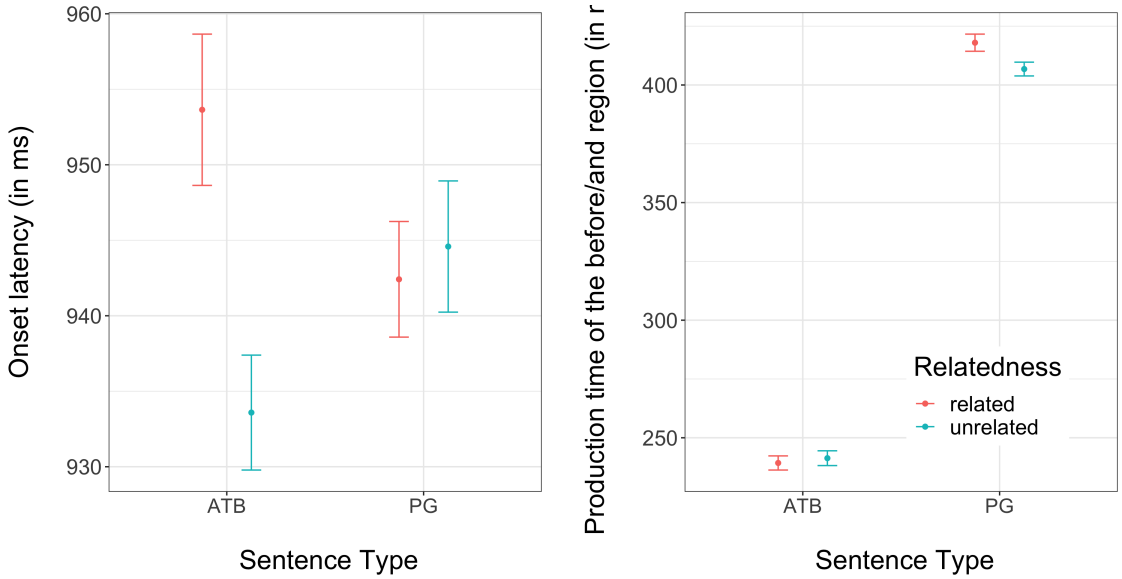
main effect of Verb Position ($\hat{\beta} = -0.002$, 95% CrI = $[-0.021, 0.016]$, $Pr(\beta > 0) = 0.584$). Planned comparisons (based on nested models) revealed evidence for an effect of Relatedness in the second verb condition ($\hat{\beta} = -0.034$, 95% CrI = $[-0.061, -0.007]$, $Pr(\beta < 0) = 0.994$), but not in the first verb condition ($\hat{\beta} = 0.006$, 95% CrI = $[-0.020, 0.032]$, $Pr(\beta < 0) = 0.329$). Thus, Experiment 1 provided evidence that the distractors we chose are effective and are specifically interfering with the verbs that were used as the second verb in Experiment 2.

Experiment 2

Overall, 35.2% of the trials (3495 out of 9920 trials) were identified as erroneous, so they were excluded from subsequent analyses. The error rates across the four conditions were as follows: ATB-Related: 35.8%; ATB-Unrelated: 36.2%; PG-Related: 36.7%; PG-Unrelated: 32.2%.

Figure 4 shows the onset latency and the production time of the *before/and* region across the four conditions. In the ATB sentences, speakers were on average 20 milliseconds slower to start speaking given the related distractors than given the unrelated distractors. In contrast, in the PG condition, speakers were if anything 3 milliseconds *faster* to start speaking given the related distractors than given the unrelated distractors. Supporting this pattern, there was evidence for a Relatedness × Sentence Type interaction ($\hat{\beta} = 0.019$, 95% CrI = $[0.004, 0.034]$, $Pr(\beta > 0) = 0.995$). Planned comparisons (based on nested models) revealed evidence for an effect of Relatedness in the ATB condition ($\hat{\beta} = -0.015$, 95% CrI = $[-0.026, -0.005]$, $Pr(\beta < 0) = 0.998$), but not in the PG condition ($\hat{\beta} = 0.004$, 95% CrI = $[-0.006, 0.014]$, $Pr(\beta < 0) = 0.226$). There was weak evidence for a main effect

Figure 4. By-subject mean onset latency (left) and by-subject mean production time of the (*before/and*) region (right) across the four conditions in Experiment 2. Error bars represent standard error of the mean.



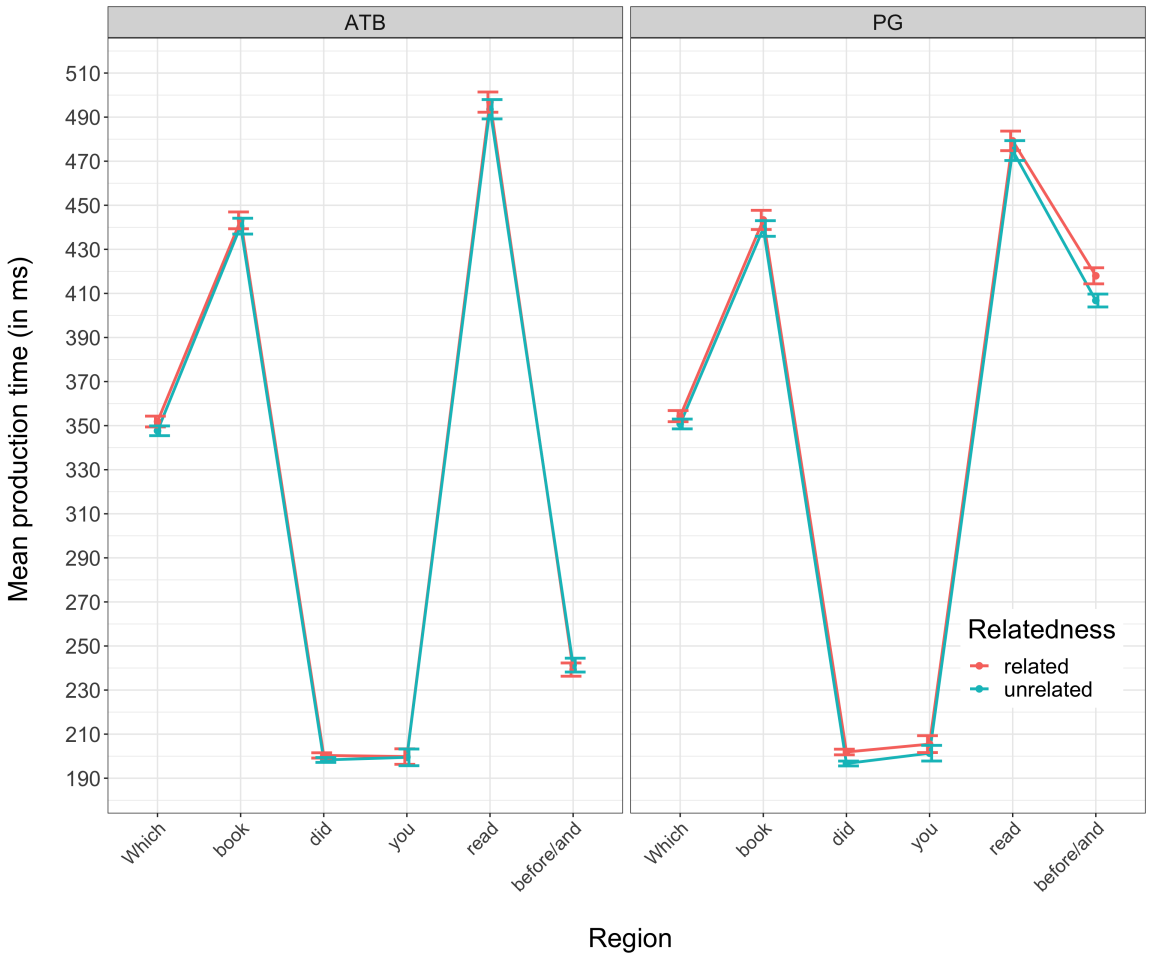
of Relatedness ($\hat{\beta} = -0.006$, 95% CrI = $[-0.013, 0.002]$, $Pr(\beta < 0) = 0.939$), but we will not interpret this effect due to the presence of the interaction involving this factor. There was no evidence for an effect of Sentence Type ($\hat{\beta} = 0.001$, 95% CrI = $[-0.007, 0.005]$, $Pr(\beta > 0) = 0.590$).

As can be seen in Figure 4, speakers took 11 milliseconds longer to speak the *before* region given related distractors than given unrelated distractors in the PG condition. In contrast, in the ATB condition, speakers took, if anything, 2 milliseconds less to speak the *and* region given related distractors than given unrelated distractors. Supporting this pattern, there was evidence for Relatedness × Sentence Type interaction ($\hat{\beta} = -0.030$, 95% CrI = $[-0.060, -0.001]$, $Pr(\beta < 0) = 0.978$). Planned comparisons (based on nested models) revealed evidence for an effect of Relatedness in the PG condition ($\hat{\beta} = -0.018$, 95% CrI = $[-0.036, -0.001]$, $Pr(\beta < 0) = 0.981$), but not in the ATB condition ($\hat{\beta} = 0.012$, 95% CrI = $[-0.015, 0.039]$, $Pr(\beta < 0) = 0.197$). There was clear evidence for a main effect of Sentence Type ($\hat{\beta} = 0.582$, 95% CrI = $[0.546, 0.619]$, $Pr(\beta > 0) = 0.999$), but this just suggests that speakers took more time to say *before* than to say *and*.

To make sure that we did not miss potential interference effects in other regions, we also conducted exploratory analyses on the production time in other regions, separately for ATB and PG sentences. Figure 5 shows the region-by-region average production time in each condition. As can be seen, production times in the ATB condition were very closely matched, and there was no evidence for an interference effect in any region in the ATB condition. In comparison, there was evidence for an interference effect in the *did* region in the PG condition ($\hat{\beta} = -0.023$, 95% CrI = $[-0.038, -0.006]$, $Pr(\beta > 0) = 0.997$). No reliable evidence for an interference effect was found in other regions.

Taken together, in the ATB condition, the verb interference effect was observed before utterance onset, but not during the utterance. In comparison, in the PG condition, the verb interference effect was observed during the utterance, specifically in the production time of the *did* and *before* regions. In the PG condition, the interference effect was numerically small in any individual region (<12ms), perhaps because the effect was distributed across more than one region. These results together suggest that speakers relatively consistently planned the second verb before speech onset in the ATB condition, while they planned the second verb later in the PG condition, though the precise timing may vary across speakers, items, and/or trials.

Figure 5. By-subject word-by-word mean production time across the four conditions in Experiment 2. Error bars represent standard error of the mean.



Discussion

We investigated the timing of verb planning in the production of sentences involving ATB and PG constructions, using the sentence-word interference task. In Experiment 1, we

established that the current task and target-distractor verb pairs can elicit verb interference effects. In Experiment 2, using the same task and the same target-distractor pairs as in Experiment 1, we showed that the semantic interference effect on verbs is observed in onset latency in ATB constructions but not in PG constructions. The opposite pattern was found in the pre-verb word production time. This pattern suggests that speakers predominantly planned the second verb before speech onset in the ATB condition, but as they articulated the sentences in the PG condition. This pattern is predicted by the syntactic account, which posits that speakers plan verbs before their syntactic complement. The results are not compatible with the conceptual account, which posits that speakers plan verbs before the patients of the event denoted by the verbs, because the filler is the patient or theme of the event denoted by the verb in both ATB and PG sentences.

Here we present a sketch of a model of structure building procedures in ATB and PG sentence production and explain how it captures the current results. In this preliminary model, structure building in the production of ATB and PG sentences proceeds as follows. Given the relevant conceptual representation, which in the current experiment was extracted from the sentences memorized, speakers first realize that the filler is associated with two events/states. Minimally, they realize that the message-level representation contains two events, that two events share the same participant, and that the shared participant corresponds to the filler. Speakers then determine whether the relationship between two events is better expressed by coordination or subordination, to choose between coordinate structures or adjunct structures. If they determined that a coordinate structure is appropriate, they construct a syntactic structure corresponding to an ATB construction, which involves two coordinated verb phrases, each with a gap. When both gaps are the direct syntactic complement of two verbs as in the current experiments, speakers plan the verbs, which are by hypothesis used to encode the syntactic dependencies between the two verbs and the filler. Because speakers plan the second verb before the articulation of the filler, they are susceptible to semantic interference on the second verb before the articulation of the filler, as the current results suggest. In comparison, if speakers determine that a subordinate structure is appropriate for expressing two events, they construct syntactic structures corresponding to PG constructions, which involves subordination. Because adverbial subordinate clauses are islands (Bondevik et al., 2020; Cattell, 1976; Chomsky, 1986; Geis, 1970; Huang, 1982; Kush et al., 2018, 2019; Sprouse et al., 2012; Stepanov, 2007), they do not posit a gap directly associated with the filler inside these subordinate clauses. Instead, they posit a pronominal element that is co-referential with the filler (i.e., a null operator or silent pronoun). This process of positing a pronominal element may occur much later than the production of the filler. Because the filler is not the syntactic complement of the verb inside the adjunct, they have no grammatical motivation to plan the second verb before beginning to speak the filler. As a consequence, when producing PG sentences, speakers are not susceptible to the semantic interference on the second verb before the production of the filler. Instead, they experience the interference effect when they actually need to plan the second verb, as the current data suggests.

This way of capturing the current results offers a unified account of the planning procedures for seemingly unrelated constructions, specifically raising-to-subject and subject control constructions. Raising-to-subject and subject control constructions show contrasts that are relevantly similar to the contrast between ATB and PG constructions. According

to some theories of raising and control (e.g., Landau, 2003; cf. Boeckx and Hornstein, 2004; Hornstein, 1999; see Polinsky, 2013 for a recent overview), when the embedded clause is in the passive voice, the matrix subject is the (underlying) syntactic complement of the embedded verb in raising-to-subject sentences, but not in subject control sentences.

In a previous ePWI study examining the timing of verb planning, Momma et al. (2020) investigated the timing of verb planning in raising-to-subject and subject control sentences like the following.

(7) a. The chef seems to be greeted by the ballerina. [raising-to-subject]
b. The chef wants to be greeted by the ballerina. [subject control]

Based on the ePWI methodology discussed in the introduction, Momma et al. found that speakers showed semantic interference on the embedded verb in onset latency in raising-to-subject sentences as in (7a), just like what the current results showed in ATB sentences. In contrast, in subject control sentences as in (7b). Instead, in subject control sentences, speakers showed verb interference later, just like what the current results showed in PG sentences.

According to some theories of grammar such as Government and Binding (Chomsky, 1981), those sentences have the following underlying representations.

(8) a. The chef$_i$ seems t$_i$ to be greeted t$_i$ by the ballerina.
b. The chef$_i$ wants PRO$_i$ to be greeted t$_i$ by the ballerina.

In (7a), the main clause subject of raising-to-subject sentences (*the chef* in the examples above) is underlyingly the syntactic complement of the embedded verb (*greeted* in the example above). This direct dependency relationship between the subject and the embedded verb is similar to the dependency involved in the ATB sentences between the filler and the second verb; that is, in both ATB and raising-to-subject sentences, there is a direct dependency between the relevant noun and the relevant verb. In comparison, in (7b), the main clause subject of the subject control sentences is the semantic object (*theme*) but not the syntactic complement of the embedded verb. That is, the relationship between the subject and the embedded verb is mediated by a null pronoun (called *PRO*), just like the relationship between the filler and the non-main verb in PG is mediated by a null operator/pronoun. Therefore, the results of the current study and the study by Momma et al. can both be accounted for by the hypothesis that speakers plan verbs before the articulation of their syntactic complements, but not before the articulation of their semantic objects that are not syntactic complements. Thus, the current approach gives us a unified account of the timing of verb planning in seemingly unrelated types of sentences.

Structure building in sentence production

The current results suggest that speakers (tend to) plan verbs before the articulation of their syntactic complements, as in the case of ATB sentences, but not necessarily before the articulation of the theme/patient of the event denoted by those verbs, as in the case of PG sentences. This contrast in turn suggests that speakers can decide whether the filler is the syntactic complement of a verb, even when the conceptual role of the filler is not sufficient to determine its grammatical status. In the current study, the conceptual role

of the filler was constant between ATB and PG sentences, so conceptual role information was not sufficient to decide whether the filler was the syntactic complement of the second verb or not. Surface phrasal structures are not sufficient either, unless they encode the distinction between the two types of dependencies involved in ATB and PG sentences. Therefore, to capture the current results, it is insufficient to posit functional structures and surface phrase-structural representations without empty elements or some notational variants, which are often the only syntactic level of representations assumed in prominent models of sentence production (e.g., Bock and Levelt, 1994; Bock et al., 1992; Branigan and Pickering, 2017; Kempen and Hoenkamp, 1987; Levelt, 1989 among many others). Some representational device that encodes information about whether the filler is the syntactic complement of a particular verb is necessary. One such representational device is empty elements like null operators or null resumptive pronouns; they ensure that the surface phrase structure carries information about whether moved elements (like fillers) are the syntactic complement of a verb or not. Certainly, there are multiple ways of implementing such a device (see below for more in-depth discussion), so we will not commit to a particular theory of syntax here. But the main point is that theories of sentence production need to enrich the phrase-structure representations they assume, incorporating some representational device that distinguishes underlying syntactic complements from non-complements, independently of their conceptual role and of surface phrase structures. This requires building syntactic representations that are richer than usually acknowledged in most major production models that explicitly discuss the nature of the syntactic representations involved (e.g., F. Ferreira, 2000; Kempen and Hoenkamp, 1987; Levelt, 1989) and their successors (e.g., Bock and Levelt, 1994; V. S. Ferreira et al., 2018).

The current results also suggest that, given two events to be expressed, speakers can determine whether they are producing coordination (in ATB sentences) or subordination (in PG sentences) early in the production process, before starting to produce the filler associated with the coordinate or subordinate structures. This is because, for speakers to make a decision about whether to plan the second verb, they must first decide whether they are producing an ATB sentence (which involves coordination) or a PG sentence (which involves subordination). Because the distinction between coordination and subordination is a hierarchical one, this suggests that speakers at least formulate a rudimentary representation of the hierarchical relationship between clauses, either at the level of semantics, syntax, or both. Of course, this does not mean that speakers *always* know the relationship between clauses, but at least in the current task, speakers must have the capacity to represent hierarchical clausal relationships well in advance, and use them to guide their sentence planning procedures (see below for more in-depth discussion on the flexibility of sentence planning). This conception of how sentence planning proceeds is compatible with a broad class of production hypotheses known as *hierarchical incrementality* (Bock et al., 2003; Konopka, 2012; Lee et al., 2013; Van de Velde et al., 2014 among others), which state that sentence planning is (or can be) guided by hierarchical representations of message or syntax, which are presumed to be encoded early on in production processes.

Representations of ATB and PG constructions

The current results can be captured naturally under the analysis that ATB and PG constructions involve two distinct types of dependency representations: direct dependency

in ATB and indirect dependency mediated by a null element in PG. In particular, the analysis that the filler is the direct complement of the second verb in ATB but not in PG (e.g., Postal, 1993) offers a straightforward explanation of the current results, when combined with the hypothesis that speakers plan verbs only before the production of their syntactic complement. This is because, under such an analysis, the filler is actually not the syntactic complement of the second verb in PG constructions unlike in ATB constructions, and thus the second verb does not have to be planned before the filler in PG sentences. Analyses where ATB and PG constructions are representationally alike (e.g., Chaves, 2012; Hornstein and Nunes, 2002; Munn, 1992; Williams, 1990) would need to explain the current pattern based on something other than differences in the types of dependencies involved in ATB and PG constructions.

It is possible to capture the current results without assuming distinct types of dependencies for ATB and PG constructions. Specifically, it may be that the filler in PG constructions is indeed the syntactic complement of the second verb (so the dependencies involved in ATB and PG constructions are the same in type), but speakers do not plan the second verb before the filler production in PG because the filler is only optionally the syntactic complement of the second verb. Unlike in ATB, the gap of the second verb in PG can be replaced with an overt pronoun, as shown in (3b). Due to this optionality, speakers may be able to postpone any commitment to representing the filler as the syntactic complement of the second verb. If so, when producing PG sentences, speakers may not have to plan the second verb before the production of a filler with an under-specified relationship to the second verb. Under this optionality account, it is not necessary to adopt an analysis where ATB and PG constructions contain distinct types of dependencies.

However, this optionality account still needs to assume some syntactic differences between ATB and PG constructions. Under the optionality account, speakers should be able to realize that the filler is only optionally the syntactic complement of the second verb. The difference between a PG and an overt pronoun is difficult to reduce to a conceptual difference, because the propositional content of a PG sentence (e.g., *which computer did you test before repairing?*) and a minimally different sentence with an overt pronoun instead of a PG (e.g., *which computer did you test before repairing it?*) are essentially the same. So the choice between a PG and a pronoun is essentially a syntactic one. This means that speakers need to represent different syntactic structures for ATB and PG constructions to know that the filler does not have to be the syntactic complement of the second verb. Thus, the optionality account presupposes some syntactic difference(s) between ATB and PG constructions.

In sum, the current results are most naturally accommodated by representational theories assuming two distinct types of dependencies for ATB and PG constructions, although the optionality account may capture the current results without assuming two distinct types of dependencies. Regardless of which account is correct, the current results are difficult to capture without assuming some syntactic difference(s) between ATB and PG constructions that are not easily reducible to conceptual differences.

Flexibility in sentence planning

Phenomenologically, it is unlikely that speakers *must* plan verbs' lemmas before they start uttering their underlying objects (Momma & Ferreira, 2019). For instance, we can

certainly name an object that happened to be in front of us in a phrasal format (e.g., *the computer*) and continue a sentence using the already-uttered phrase as the underlying object (e.g., *the computer fell from the table*). In fact, if we were incapable of doing so, it would be impossible to construct object-initial, verb-final structures in the well known cloze task (Taylor, 1953), contrary to fact. Consistent with this intuition, various experimental studies also suggest that planning scope is flexible, at least in terms of noun phrase planning without direct dependency relationships (F. Ferreira and Swets, 2002; Konopka, 2012; Konopka and Meyer, 2014; Wagner et al., 2010). In addition, a previous corpus study suggests that speakers may plan verbs' lemmas sometimes early and sometimes late (Van de Velde et al., 2015). This study showed that, in Dutch, verb bias on structural choice in the dative alternation has a weaker (though still significant) effect in verb-final clauses than in verb-initial clauses. Another study investigating speech errors involving case markers in Japanese (Iwasaki, 2010) showed that, at least occasionally, speakers may exploit the correlation between theme/patient thematic roles and accusative case marking to (wrongly) choose the accusative case marker for the subjects of unaccusative or passive verbs, which require nominative case marking. This suggests that Japanese speakers may not always use verbs' lexical information to select a case marking. Thus, it is unlikely that verbs, or the event concepts associated with them, are obligatorily planned before the utterance of their syntactic object (see F. Ferreira and Swets, 2002; Konopka, 2012; Levelt, 1989; Wagner et al., 2010 for more general discussion on the flexibility of planning scope that is not restricted to the planning of verbs).

There are two ways to capture this flexibility in verb planning. One approach is to accept that speakers have two ways of grammatically encoding underlying objects (verb-dependent and verb-independent encoding), and speakers choose between these two strategies depending on the circumstances. Under this approach, the generalization about the timing of verb planning reflects speakers' *tendency* to plan verbs before a certain point in time. This approach is appealing as it has broader empirical coverage, but it is at the same time not constraining. An alternative approach is to assume that verbs' lemmas *must* be planned before the object argument can be integrated into the overall structure of a sentence, rather than before the object argument is articulated. That is, articulation might precede grammatical integration, and only the process of grammatical integration requires verbs' lemmas. Normally, speakers grammatically integrate a phrase to be produced before articulation, but this may not be a strict requirement of the production system. This approach is more constraining than the first approach, but the challenge for this approach is that it is not easy to know when speakers utter noun phrases in isolation vs. as an integrated part of a sentence. So, unfortunately, we are not able to support or reject either approach in the current study.

Related to the flexibility issue, we do not argue that the current results necessarily hold for everyday speaking. For instance, the current study does not necessarily suggest that speakers always plan verbs' lemmas before speaking ATB sentences or that speakers do not plan verbs' lemma before speaking PG sentences. Given that the current study is based on a sentence-recall task, it is likely that speakers have more certainty about what structures to use prior to speech planning than in everyday speaking. This might have magnified the difference in the timing of verb planning between ATB and PG sentence production. However, this does not constitute a basis for disregarding the experimental

results (Mook, 1983). The primary purpose of the current experiment (and many other psychological experiments) is not to make predictions about how speakers should behave in everyday life, but to test theories about speakers' cognitive capacities. Even if the current task does not resemble real-world situations, the difference in the timing of verb planning between ATB and PG production demands an explanation. Our explanation is that the sentence-initial object fillers are directly syntactically dependent on verbs in ATB sentences but not PG sentences, and consequently, at least when speakers can represent this difference (as in the current task), speakers plan verbs before the sentence-initial object fillers in ATB but not in PG to establish the syntactic dependency.

Memory of sentences

In the memory literature, it has been argued that memory of sentences, even in short-term recall tasks, is not verbatim memory (Potter and Lombardi, 1990, 1998; Sachs, 1974). When people are able to recall a sentence they memorized accurately, this successful recall is due to the regeneration of the sentence structure from conceptual memory using the unordered set of recently activated words. Put differently, sentence recall involves grammatical encoding in sentence production (see Bock, 1982; V. S. Ferreira, 2003 for a similar view). The current results support this hypothesis about how people recall sentences for production. If sentences are recalled from verbatim memory, it is unclear how the time-course of sentence planning could differ between ATB and PG sentences as the current results suggest. Thus, the current results indirectly support the basic claims of the regeneration hypothesis (Potter and Lombardi, 1990).

If sentence recall reflects sentence production processes, as the current results suggest, it is good news for researchers who study sentence production. One of the primary methodological challenges for studying sentence-level production is that it is hard to elicit grammatically complex sentences, which are critical in investigating the fine details of mental representations of sentences in production. Previous studies (e.g., Chang et al., 2003; V. S. Ferreira, 2003; McDonald et al., 1993; Potter and Lombardi, 1998) have shown that various properties of sentence production mechanisms can be studied using sentence recall, and the current study adds to those previous studies by showing that, using a variant of sentence recall tasks we call the sentence-word interference task, it is possible to gain insights about not only factors influencing the final form of sentences, but also the time-course of the planning processes involved in grammatically complex sentences. We thus hope that the current study contributes to establishing a methodological basis for studying various types of sentences that have been important in studying the structures of sentences, to facilitate cross-talk between theories of sentence production and theories of grammatical representations.

Conclusion

The current study examined the timing of verb planning in ATB and PG sentence production, to better understand how speakers plan grammatically complex sentences involving multiple dependencies. The results suggest that speakers predominantly plan both verbs before starting to speak the filler associated with them in ATB sentences, but not in PG sentences. This timing contrast is predicted by the combination of the production hy-

788 pothesis that speakers plan verbs before the production of their syntactic complement and
789 the representational hypothesis that assumes distinct types of dependencies in ATB and
790 PG constructions. Thus, the current study supports these hypotheses about production
791 processes and syntactic representations. More generally, these hypotheses in turn suggest
792 that sentence production processes are guided by rich syntactic representations that are not
793 reducible to conceptual representations.

References

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456.

Baker, M. C. (2009). Language universals: Abstract but not mythological. *Behavioral and Brain Sciences*, *32*(5), 448.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278.

Bock, K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*(1), 1.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. (2003). Minding the clock. *Journal of Memory and Language*, *48*(4), 653–685.

Bock, K., & Levelt, W. J. (1994). Language production: Grammatical encoding. *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological review*, *99*(1), 150.

Boeckx, C., & Hornstein, N. (2004). Movement under control. *Linguistic inquiry*, *35*(3), 431–452.

Bondevik, I., Kush, D., & Lohndal, T. (2020). Variation in adjunct islands: The case of norwegian. *Nordic Journal of Linguistics*, *doi:10.1017/S0332586520000207*, 1–32.

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, *40*.

Browning, M. (1982). *Null operator constructions* (Doctoral dissertation). Massachusetts Institute of Technology.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of english and spanish. *Cognition*, *109*(2), 274–280.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, *54*(4), 592–609.

Bruening, B., & Al Khalaf, E. (2017). *Reconstruction and linear order in atb movement and parasitic gap constructions* [Manuscript, University of Delaware and University of Jordan].

Bürki, A., Alario, F., Vasishth, S., et al. (2020). When words collide: Bayesian meta-analyses of distractor and target properties in the picture-word interference paradigm. *arXiv preprint arXiv:2008.03972*.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture-word interference paradigm: Does the response set matter? *Cognition*, *75*(2), B51–B64.

Cattell, R. (1976). Constraints on movement rules. *Language*, *52*(1), 18–50.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29–49.

Chaves, R. P. (2012). On the grammar of extraction and coordination. *Natural Language & Linguistic Theory*, *30*(2), 465–512.

Chomsky, N. (1981). *Lectures on government and binding: The pisa lectures*. Foris Publications.

Chomsky, N. (1986). *Barriers* (Vol. 13). MIT press.

Contreras, H. (1984). A note on parasitic gaps. *Linguistic inquiry*, *15*(4), 698–701.

Contreras, H. (1993). On null operator structures. *Natural Language and Linguistic Theory*, *11*(1), 1–30.

Culicover, P. (2001). Parasitic gaps: A history. In P. Culicover & P. Postal (Eds.), *Parasitic gaps* (pp. 3–68). MIT Press.

Culicover, P., & Postal, P. M. (2001). *Parasitic gaps* (Vol. 35). MIT Press.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, *10*(1), 111–117.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*(2), 195–209.

De Smedt, K. (1990). Ipf: An incremental parallel formulator. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 167–192). Academic Press.

De Vries, M. (2017). Across-the-board phenomena. In M. Everaert & H. Riemsdijk (Eds.), *The wiley blackwell companion to syntax, second edition* (pp. 1–31). Wiley Online Library.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, *93*(3), 283.

Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and cognitive processes*, *23*(4), 583–608.

Engdahl, E. (1983). Parasitic gaps. *Linguistics and philosophy*, *6*(1), 5–34.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). Psychology press.

Ferreira, F., & Swets, B. (2002). How incremental is language production? evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84.

Ferreira, V. S., Morgan, A., & Slevc, L. (2018). Grammatical encoding. *The oxford handbook of psycholinguistics, 2nd ed.* (pp. 432–457). Oxford University Press.

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language*, *48*(2), 379–398.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology*, *40*(4), 296–340.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of learning and motivation* (pp. 133–177). Academic Press.

Geis, M. (1970). *Adverbial subordinate clauses in English* (Doctoral dissertation). Massachusetts Institute of Technology.

Goldsmith, J. (1985). A principled exception to the Coordinate Structure Constraint. In W. H. Eilfort, P. D. Kroeber, & K. L. Peterson (Eds.), *Proceedings of the chicago linguistic society 21* (pp. 133–143).

Griffin, Z. M., & Ferreira, V. S. (2006). Properties of spoken language production. In M. Gernsbacher & M. Traxler (Eds.), *Handbook of psycholinguistics (second edition)* (pp. 21–59). Academic Press.

Haik, I. (1985). *The syntax of operators* (Doctoral dissertation). Massachusetts Institute of Technology.

Harley, H., & Stone, M. (1992). The 'no agent idioms' hypothesis. In R. Folli, C. Sevdali, & R. Truswell (Eds.), *Syntax and its limits* (pp. 251–276). Oxford University Press.

Hornstein, N. (1999). Movement and control. *Linguistic inquiry*, *30*(1), 69–96.

Hornstein, N., & Nunes, J. (2002). On asymmetries between parasitic gap and across-the-board constructions. *Syntax*, *5*(1), 26–54.

Huang, C.-T. J. (1982). *Logical relations in Chinese and the theory of grammar* (Doctoral dissertation). Massachusetts Institute of Technology.

Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in english and korean. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1363.

Iwasaki, N. (2010). Incremental sentence production: Observations from elicited speech errors in japanese. *Processing and producing head-final structures* (pp. 131–151). Springer.

Jescheniak, J. D., Wöhner, S., Bethcke, H. S., & Beaupain, M. C. (2020). Semantic interference in the picture-word interference task: Is there a pre-lexical, conceptual contribution to the effect? *Psychonomic Bulletin & Review*, 1–6.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. CSLI Publications.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive science*, *11*(2), 201–258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*(2), 185–209.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, *66*(1), 143–162.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40.

Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck & L. Zarin (Eds.), *Phrase structure and the lexicon* (pp. 109–137). Springer.

Kratzer, A. (2003). The event argument and the semantics of verbs. *Ms., University of Massachusetts, Amherst*.

Kush, D., Lohndal, T., & Sprouse, J. (2018). Investigating variation in island effects. *Natural Language and Linguistic Theory*, *36*, 743–779.

Kush, D., Lohndal, T., & Sprouse, J. (2019). On the island sensitivity of topicalization in norwegian: An experimental investigation. *Language*, *95*, 393–420.

Landau, I. (2003). Movement out of control. *Linguistic inquiry*, *34*(3), 471–498.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Lee, E.-K., Brown-Schmidt, S., & Watson, D. G. (2013). Ways of looking ahead: Hierarchical planning in language production. *Cognition*, *129*(3), 544–562.

Levelt, W. J. (1989). *Speaking: From intention to articulation* (Vol. 1). MIT press.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*(1), 1–38.

Levine, R. D., & Hukari, T. E. (2009). The unity of unbounded dependency constructions. *Journal of Linguistics*, *45*(1), 232.

Levine, R. D., & Sag, I. (2003). Some empirical issues in the grammar of extraction. In S. Muller (Ed.). CSLI Publications.

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of multivariate analysis*, *100*(9), 1989–2001.

Lindsley, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, *7*(1), 1–19.

Lindsley, J. R. (1976). Producing simple utterances: Details of the planning process. *Journal of Psycholinguistic Research*, *5*(4), 331–354.

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, *31*(6), 713–733.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition*, *7*(6), 485–495.

Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture–word interference. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(4), 269.

Marantz, A. (1981). *On the nature of grammatical relations* (Doctoral dissertation). Massachusetts Institute of Technology.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive psychology*, *25*(2), 188–230.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of memory and Language*, *35*(4), 477–496.

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 907.

Momma, S. (2016). *Parsing, generation and grammar* (Doctoral dissertation).

Momma, S., & Ferreira, V. S. (2019). Beyond linear order: The role of argument structure in speaking. *Cognitive psychology*, *114*, 101228.

Momma, S., Slevc, L. R., & Phillips, C. (2016). The timing of verb selection in japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 813.

Momma, S., Slevc, L. R., & Phillips, C. (2018). Unaccusativity in sentence production. *Linguistic Inquiry*, *49*(1), 181–194.

Momma, S., Wilson, M., & Ferreira, V. (2020). *Syntax guides sentence production* [the 33rd CUNY conference on human sentence processing].

Mook, D. G. (1983). In defense of external invalidity. *American psychologist*, *38*(4), 379.

Munn, A. (1992). A null operator analysis of atb gaps. *The Linguistic Review*, *9*(1), 1–26.

Nissenbaum, J. W. (2000). *Investigations of covert phrase movement* (Doctoral dissertation). Massachusetts Institute of Technology.

Ogihara, T. (1994). Adverbs of quantification and sequence-of-tense phenomena. *Semantics and Linguistic Theory*, *4*, 251–267.

Polinsky, M. (2013). Raising and control. In M. den Dikken (Ed.), *The cambridge handbook of generative syntax*. Cambridge University Press.

Postal, P. M. (1993). Parasitic gaps and the across-the-board phenomenon. *Linguistic Inquiry*, *24*(4), 735–754.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, *29*(6), 633–654.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language*, *38*(3), 265–282.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107–142.

Ross, J. (1967). *Constraints on variables in syntax.* (Doctoral dissertation). Yale University.

Rummer, R., Schweppe, J., & Martin, R. C. (2013). Two modality effects in verbal short-term memory: Evidence from sentence recall. *Journal of Cognitive Psychology*, *25*(3), 231–247.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, *2*(1), 95–100.

Sag, I. A. (1983). On parasitic gaps. *Linguistics and Philosophy*, *6*(1), 35–45.

Sauppe, S. (2017). Word order and voice influence the timing of verb planning in german sentence production. *Frontiers in psychology*, *8*, 1648.

Schmerling, S. F. (1972). Apparent counterexamples to the coordinate structure constraint. *Studies in the Linguistic Sciences*, *2*(1), 91–104.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of memory and language*, *29*(1), 86–102.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in german. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1368.

Schriefers, H., Teruel, E., & Meinshausen, R.-M. (1998). Producing simple sentences: Results from picture–word interference experiments. *Journal of Memory and Language*, *39*(4), 609–632.

Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language*, *1*, 82–123.

Stepanov, A. (2007). The end of CED? minimalism and extraction domains. *Syntax*, *10*(1), 80–126.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism quarterly*, *30*(4), 415–433.

Van de Velde, M., Kempen, G., & Harbusch, K. (2015). Dative alternation and planning scope in spoken language: A corpus study on effects of verb bias in vo and ov clauses of dutch. *Lingua*, *165*, 92–108.

Van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing "easy" and "hard" events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*(1), 124–144.

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of phonetics*, *71*, 147–161.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 423.

Williams, E. (1978). Across-the-board rule application. *Linguistic Inquiry*, *9*(1), 31–43.

Williams, E. (1990). The atb theory of parasitic gaps. *The Linguistic Review*, *6*(2), 265–279.

Williams, E. (1992). Adjunct control. In R. K. Larson, S. Iatridou, U. Lahiri, & J. Higginbotham (Eds.), *Control and grammar(48)* (pp. 297–322). Kluwer Academic Publishers.

Zehr, J., & Schwarz, F. (2018). Penncontroller for internet based experiments (ibex).

Appendix

Stimulus list

Target sentences used in Experiment 2 (Across-the-board/Parasitic gap). The words inside the parentheses are distractor words (Related/Unrelated).

1. Which apple did you pick and pack/before packing? (fill/assist)
2. Which article did you read and criticize/before criticizing? (recommend/mix)
3. Which artist did you meet and support/before supporting? (adopt/cook)
4. Which book did you read and burn/before burning? (melt/accept)
5. Which box did you assemble and fill/before filling? (pack/solve)
6. Which box did you close and hide/before hiding? (open/revise)
7. Which box did you organize and close/before closing? (fold/recommend)
8. Which boy did you bark at and bite/before biting? (squeeze/submit)
9. Which bread did you cut and bake/before baking? (cook/advertise)
10. Which cake did you frost and eat/before eating? (drink/publish)
11. Which candidate did you meet and hire/before hiring? (rent/grade)

12. Which car did you fix and sell/before selling? (buy/fold)
13. Which cat did you name and adopt/before adopting? (support/water)
14. Which chair did you adjust and paint/before painting? (decorate/drink)
15. Which cheese did you cut and melt/before melting? (burn/understand)
16. Which computer did you test and repair/before repairing? (help/stir)
17. Which criminal did you shout at and arrest/before arresting? (suspend/pet)
18. Which customer did you appease and assist/before assisting? (serve/pack)
19. Which diamond did you examine and/before buying? (sell/dismiss)
20. Which dog did you play with and feed/before feeding? (water/announce)
21. Which donkey did you clean and hug/before hugging? (pet/classify)
22. Which donkey did you wash and ride/before riding? (climb/propose)
23. Which door did you knock on and open/before opening? (hide/reject)
24. Which draft did you reread and edit/before editing? (revise/help)
25. Which email did you read and forward/before forwarding? (release/explain)
26. Which employee did you investigate and suspend/before suspending? (arrest/describe)
27. Which flowers did you smell and water/before watering? (feed/adopt)
28. Which formula did you memorize and understand/before understanding? (forget/melt)
29. Which fruits did you peel and mix/before mixing? (stir/criticize)
30. Which girl did you praise and help/before helping? (repair/edit)
31. Which guitar did you customize and smash/before smashing? (break/release)
32. Which horse did you feed and pet/before petting? (hug/arrest)
33. Which house did you renovate and advertise/before advertising? (exhibit/bake)
34. Which idea did you elaborate and present/before presenting? (announce/climb)
35. Which juice did you pasteurize and drink/before drinking? (eat/paint)
36. Which lemon did you wash and squeeze/before squeezing? (bite/approve)
37. Which manuscript did you proof-read and submit/before submitting? (propose/bite)

38. Which manuscript did you review and accept/before accepting? (praise/burn)
39. Which mountain did you read about and climb/before climbing? (ride/present)
40. Which movie did you watch and explain/before explaining? (describe/forward)
41. Which movie did you watch and praise/before praising? (accept/break)
42. Which paper did you check and revise/before revising? (edit/hide)
43. Which phone number did you write down and forget/before forgetting? (understand/exhibit)
44. Which plan did you modify and announce/before announcing? (present/repair)
45. Which plant did you dissect and classify/before classifying? (grade/hug)
46. Which policy did you refine and propose/before proposing? (submit/ride)
47. Which problem did you simplify and solve/before solving? (answer/fill)
48. Which proposal did you read and reject/before rejecting? (dismiss/open)
49. Which pumpkin did you carve and decorate/before decorating? (paint/rent)
50. Which question did you clarify and answer/before answering? (solve/serve)
51. Which request did you read and dismiss/before dismissing? (reject/buy)
52. Which room did you clean and rent/before renting? (hire/decorate)
53. Which sculpture did you paint and exhibit/ before exhibiting? (advertise/forget)
54. Which shirt did you wash and fold/before folding? (close/sell)
55. Which song did you edit and release/before releasing? (forward/smash)
56. Which soup did you warm up and stir/before stirring? (mix/feed)
57. Which story did you modify and publish/before publishing? (approve/eat)
58. Which student did you interview and recommend/before recommending? (criticize/close)
59. Which tea did you sweeten and serve/before serving? (assist/answer)
60. Which vaccine did you test and approve/before approving? (publish/squeeze)
61. Which vase did you polish and break/before breaking? (smash/praise)
62. Which vegetable did you wash and cook/before cooking? (bake/support)
63. Which wine did you sip and grade/before grading? (classify/hire)
64. Which word did you memorize and describe/before describing? (explain/suspend)